Module Project Template

Every group submits one project proposal. The recommended length is 1000-1500 words (including the template).

Group number and group members: **Group 2:** Aidan Birdi, Zijia (Anthony) Lin, Muzammil Farhan, Muyi Su
Group advisor: Emily Muller
Title: Generation and Evaluation of Clinically Realistic Images using Generative Adversarial Networks
Research question: Can a GAN-based model trained on medical MRI brain images consistently generate novel and clinically realistic images? We will evaluate our generated images using quantitative and qualitative approaches, and ask radiologists to evaluate and opine on the generated data and controls. Reach: How might traversing the latent space of our model be clinically interpreted?
Background: In order to realise the potential for machine learning in healthcare, models need to be trained on a lot of data providing good overall generalisation. Medical data is scarce, due to privacy concerns as well as the cost of medical examination. One potential means of increasing the amount of data is synthetic data generation [Kim, S., Kim, B. & Park, H. (2021) Synthesis of brain tumor multicontrast MR images for improved data augmentation. *Medical Physics (Lancaster).* 48 (5), 2185–2198. 10.1002/mp.14701.]. In this research, we will develop image-based machine learning models for predicting neurological illnesses present in MRI's. If diagnostic photographs are to be utilised in a publication or placed into the public domain, patient consent may be required (Clinical Practice Com-mittee, 2000). Synthetic data generation aids in overcoming the privacy concerns associated with diagnostic medical picture data and addressing the deficiency of positive cases for each pathology [Kim, S., Kim, B. & Park, H. (2021) Synthesis of brain tumor multicontrast MR images for improved data augmentation. *Medical Physics (Lancaster).* 48 (5), 2185–2198. 10.1002/mp.14701.]. GANs are commonly used in medical picture synthesis [Kim, S., Kim, B. & Park, H. (2021) Synthesis of brain tumor multicontrast MR images for

improved data augmentation. *Medical Physics (Lancaster).* 48 (5), 2185–2198. 10.1002/mp.14701.]. Another obstacle to the implementation of supervised training methods is the lack of medical image annotation specialists. Although there are ongoing collaborative efforts across multiple healthcare agencies to build large open access datasets, e.g. Biobank, the National Biomedical Imaging Archive (NBIA), The Cancer Imaging Archive (TCIA), and the Radiologist Society of North America (RSNA), this issue persists and limits the number of images that researchers can access. Scaling, rotation, inversion, translation, and elastic deformation are typical techniques for enhancing training sample (Simard et al., 2003). However, these changes do not account for variances caused by different imaging protocols or sequences, much alone variations in the size, shape, location, and appearance of particular disease. GANs offer a more general answer and have been utilised in various works with promising results for enhancing training photos. [Yi, X., Walia, E. & Babyn, P. (2019) Generative adversarial network in medical imaging: A review. *Medical Image Analysis.* 58 101552. 10.1016/j.media.2019.101552.]

Applications of Generative Adversarial Networks (GANs), a generative model based on game theory, have made significant progress. Due to the power of competitive training and deep neural networks, GANs are capable of generating realistic images and have demonstrated significant progress in numerous image production and editing models. Goodfellowet al. (2014) suggested GANs as a novel technique to train a generative model. Typically, GANs are deployed in a semi-supervised environment. They are comprised of two adversarial models: a generating model G that captures the data distribution and a discriminative model D that assesses the probability that a sample comes from the training data as opposed to G. G can only learn through interaction with D. (G has no direct access to real images). D, in comparison, has access to both synthetic and authentic samples. They do not directly simulate the probability distribution that creates the training data, in contrast to FVBNs (Fully Visible Belief Networks) and VAE (Variable Autoencoder). In reality, G transfers the noise vector z in the latent space to an image, while D is defined as classifying an input as a true image (near to 1) or a fake image (close to 0). Typically, images generated by GANs are less blurry and more realistic than those generated by other generative models. This is why we employ this adaptive approach for our inquiry to synthesise images. [Nasr Esfahani, S. & Latifi, S. (2019) Image Generation with Gans-based Techniques: A Survey. *International Journal of Computer Science and Information Technology.* 11 (5), 33–50. 10.5121/ijcsit.2019.11503.]

Interdisciplinary aspects of the project:
(if some of the project's aspect span disciplines, let us know)

1. Machine learning in medicine is a strongly growing field, and many machine learning algorithms are based on physical models such as diffusion model and Poisson flow model. The application of machine learning draws from computer science, maths and statistics, which we will be applying to a medical context. Medical image interpretation is very complicated and requires years of experience

as a medical practitioner to be clinically safe in requesting, interpreting and acting upon medical images.

2. MRI scans are based on Physical principles. Interpretation of MRI scans requires specialist medical knowledge and interpretation.

3. Survey (Likert scales) - statistical analysis?

Breakdown to individual steps:
(each step should contain a reasonable level of detail)

1. [Muzammil] Find relevant datasets, using medical knowledge and data science principles
2. [Zijia] Sort through dataset to find homogenous images (this could be complicated and involve use of semi-supervised learning, if there is a lot of images to sort through – we may for example need to use an ML model which can detect images in a specific medical plane (coronal, sagittal or axial))
3. [week1] Write a GAN model
4. [week2] Write questionnaire for evaluation of GAN (e.g. To what extent does this look like a real image etc.)
5. [week2] Test GAN on dummy images
6. Test GAN on a limited proportion of dataset
7. Tweak GAN for best results (while keeping in mind, any likelihood of overfitting – use different parts of the same dataset)
8. [week3] Apply GAN to whole dataset, using Imperial GPU HPC
9. [week3] Generate an appropriate number of images.
10. [week4] Quantitively appraise GAN using generative model metrics (e.g. Sørensen–Dice coefficient)
11. [week4] Send questionnaire to clinical radiologists and use qualitative research methodology to write-up our findings
12. Create academic poster on our topic
13. Extensions

Define the project core (aka the minimum viable product):
(this is the minimal version of your project that you should develop first; defining the core will help ensure that you deliver a working product)

Out MVP will be a simple GAN model applied to medical images.

Define the project extensions:
(this are additional features that you will implement after the core of the project is finished)

We could apply different generative models to our dataset, to compare. These may include:

- Variational autoencoders
- Diffusion Models?

- Poisson Flow generative models?

We could work on methods of quantitatively appraising our output, such as using the Sørensen–Dice coefficient.

Software tools needed for the projects:
(new code, libraries or packages, existing scientific software, …)

We will use Python as the programming language in this project. Regular packages like Pandas, NumPy, SciPy, and matplotlib will be used for data processing and visualization. We will also use the GPU TensorFlow package for machine learning.

Datasets that will be used in the project:

Data set	Diseases	# Patients	# Images	Plane	Format	URL	Notes	
Brain tumor data set	Glioma …							
		253			.jpg/ .jpeg	https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection		
			3060			https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection?select=no		
BRAINS	N/A	808				https://www.sciencedirect.com/science/article/pii/S1053811916000331	808 healthy volunteers	
UCSF-PDGM	Gliomas Control?	501	11,523		NIfTI format , 156 GB	https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=119705830		

*** https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset**
https://figshare.com/articles/dataset/brain_tumor_dataset/1512427

(include a short statement about licensing)

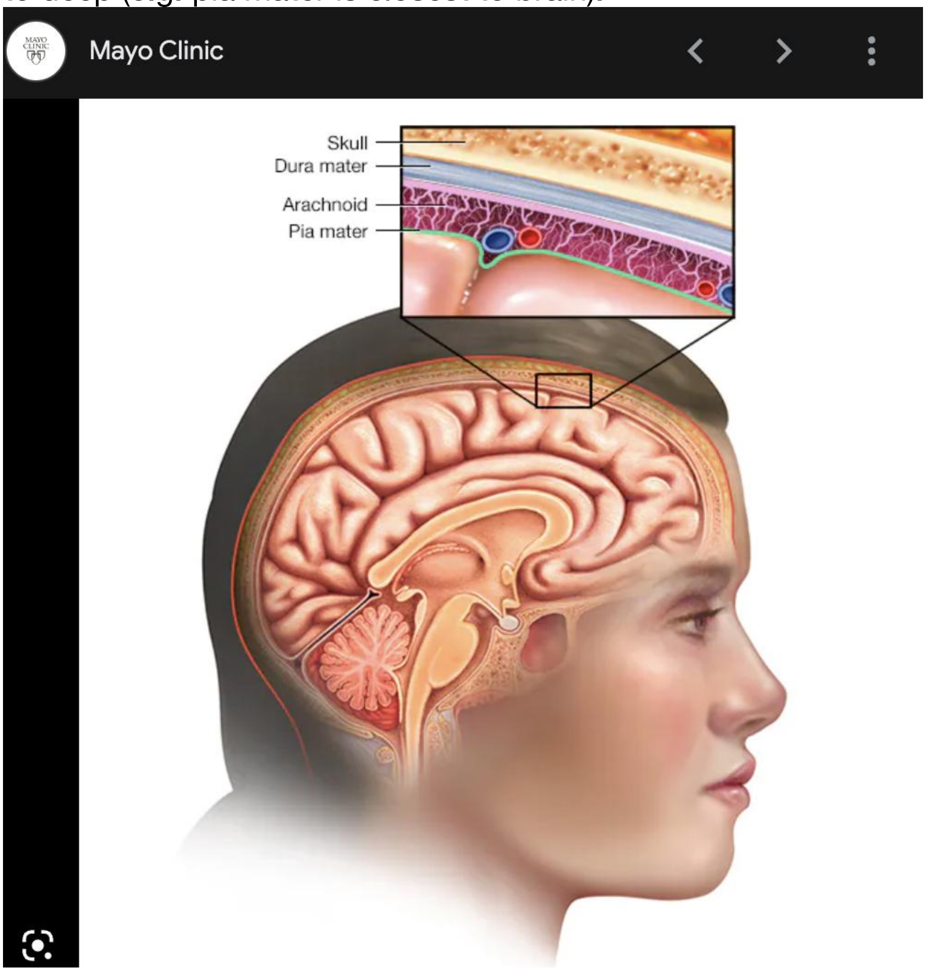
Systematically, and medically informed, find datasets linked to datasets

Licensing: This dataset is under the Creative Commons License which states that we are free to share and adapt the material for any purpose. That means we can remix, transform and build upon the material for even commerical purposes. We must however, give appropriate credit and provide a link to the license. We should also indicate if any changes were made.

The data is made up of 3064 T1 weighted contrast-enhanced images from 233 patients with 3 kinds of brain tumour.

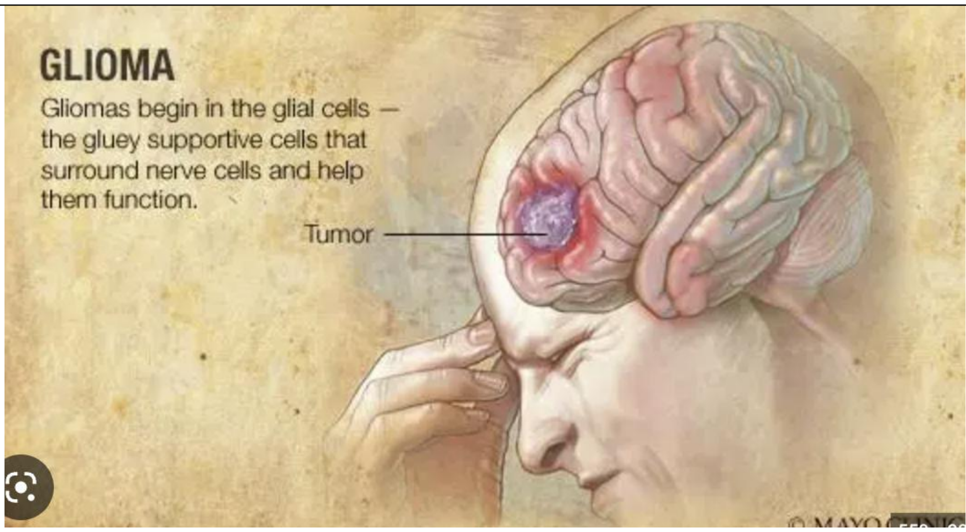
Meningioma tumour: 708/3064
Glioma tumour: 1436/3064
Pituitary tumour: 930/3064

Medical background of different tumour types:
Meningioma:

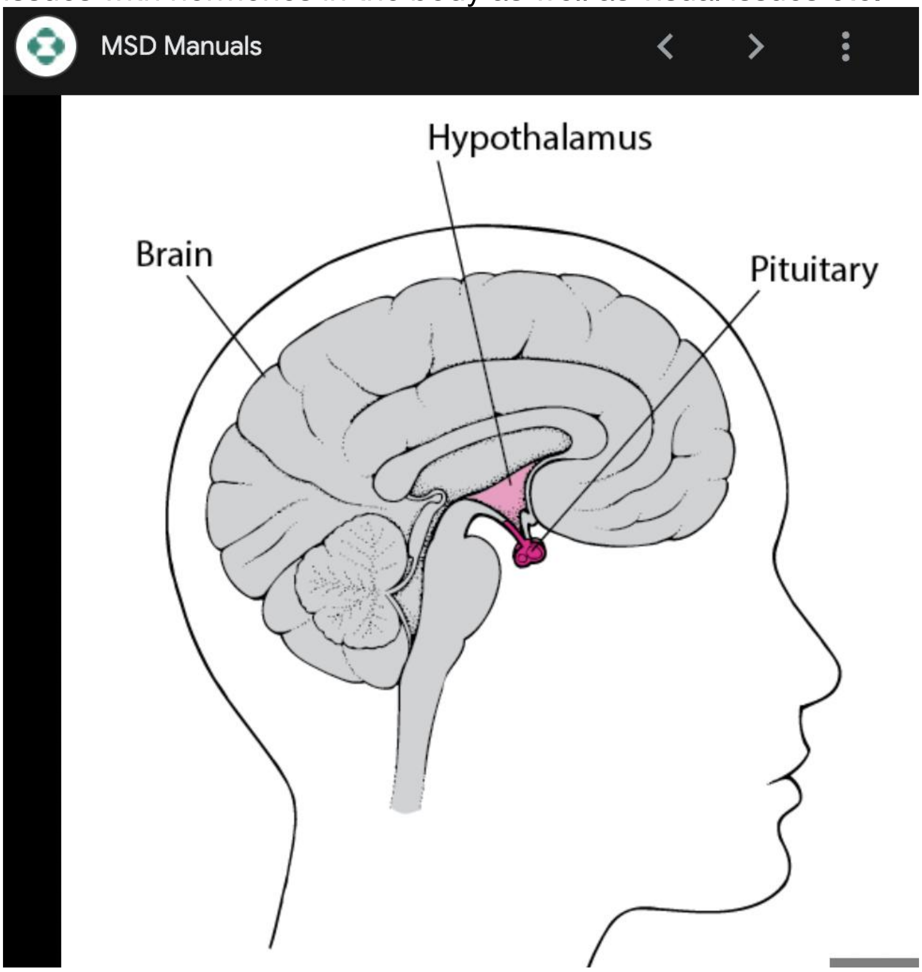
- These are tumours that occur on essentially the outermost layer of tissue that surrounds the brain. The meninges is the outermost layer made up of 3 layers – Dura mater, Arachnoid mater and Pia Mater ordered from most superficial to deep (e.g. pia mater is closest to brain).
- 

Glioma:

- Tumour of glial cells which are essentially non-neuronal cells. Glial cells surround neuronal cells and help them function.

- 

Pituitary Tumour:

- The pituitary gland is a pea-sized gland located at the base of your brain. It is often known as the 'master gland' as it sends many hormones that instruct other glands to produce hormones. It is very important clinically as well as the location being near optic nerve fibres as well as near the cavernous sinus where many other nerves travel through. A tumour here can result in many issues with hormones in the body as well as visual issues etc.
- 

Our Preferred Choice of Tumour MRI:

- I believe for our research purposes it would be better to use pituitary tumour MRIs as the location of the pituitary tumour would be relatively stable and so would provide more homogenous data. That would mean 930 MRI slides we can take advantage of. This may be changed depending on the model and the type of data that would best suit it.

Prerequisites to Data:

- There are virtually no prerequisites to this data, it can be downloaded from the link pasted.

Format of Data:

- The data is organised in 4 .zip files containing 766 slices each. The data is organised in Matlab format (.mat).

```
cjdata.label: 1 for meningioma, 2 for glioma, 3 for pituitary tumor
cjdata.PID: patient ID
cjdata.image: image data
cjdata.tumorBorder: a vector storing the coordinates of discrete points on tumor border.
                For example, [x1, y1, x2, y2,...] in which x1, y1 are planar coordinates on tumor
border.
                It was generated by manually delineating the tumor border. So we can use it to
generate
                binary image of tumor mask.
cjdata.tumorMask: a binary image with 1s indicating tumor region
```

-
- This is a screenshot taken from the link pasted above that outlines how the data is organised.
- **I believe many of the images are in the transverse plane so we could build a model that uses images taken in this plane however this would have to be double checked.**
- **My one concern is that this dataset may not contain normal MRIs so we may have to gather this from another data set.**

Required hardware:
(let us know if your laptops will be sufficient or if you require access to a Linux server)

This project will be largely machine learning based. Training machine learning models are computationally intensive, and they are made up of various matrix multiplications [1]. A GPU is preferred to speed up the process. GPUs have high level parallelism, so they can perform almost all the tasks at the same time [1], thus having very good performances. As we are processing medical images, a GPU is almost required [1]. Currently, Zijia (Anthony) has a laptop with RTX 3070 Laptop GPU and Tensor flow installed, so we might be able to use Zijia's laptop to run test GANs. Aidan has a desktop with an RTX 3080 GPU with Tensorflow installed, as a backup.

We could also use Google Colab GPU to run test samples, although more research needs to be done on the type of data that we can safely put through Google Colab. In Aidan's preliminary testing, Google Colab GPU was only about 20% slower than the desktop RTX 3080 at training the Tensorflow tutorial GAN.

We will likely need to use GPU HPC to run the larger model, once we have refined it

– as the dataset is quite large.
Agreed contributions from each group member: (for each group member, specify agreed contributions – some level of detail is required, remember that the project has multiple aspects – research, programming, GitHub repository, presentation, coordination…) Muzammil and Aidan are responsible for retrieving datasets to be used for GAN training. They will use medical expertise to justify what datasets are reasonable. Muzammil and Muyi will be responsible for data safety and ethics. Muzammil and Muyi will collaborate to import the datasets and produce visualisation algorithms for the data and the results. Anthony and Muyi (with possible need of medical expertise from Aidan or Muzammil) will work on homogenising the datasets to make them useful for the algorithm. It is difficult to know whether this step could be larger than expected until we start the project. It may be that we need to employ other ML techniques such as semi-supervised learning to help with homogenising data if there is a lot of it. More team involvement may be needed. Aidan and Anthony will cooperate to carry out the machine learning tasks in the project. This will mainly involve tweaking existing models so that it is suited for the type of data and training that we need. They will train the basic GAN using Anthony's laptop, Aidan's desktop or Google Colab if possible. Then they will use GPU HPC for the full dataset. Medical students—Aidan and Muzammil—will be carrying out medicine-related research, while other research will be carried out by Anthony and Muyi. Aidan is responsible for keeping the GitHub repository All members will cooperate on the presentation and the poster.
Agreed knowledge sharing: (it is strategic to divide work between members, however, for everyone to benefit fully you should also share work between group members – for each group member, specify what and how they would like to learn from others and teach others) Since Aidan has the most knowledge of machine learning, he will guide other members through the machine learning tasks in the project. Muzammil is new to Python programming. Every other member of the group will help Muzammil with basic Python programming. Aidan and Muzammil will be sharing essential medical knowledge for the project to Anthony and Muyi during the project. Anthony and Muyi will share essential physical knowledge with Aidan and Muzammil.
Agreed timeline: (to best of your abilities, estimate the rate of your progress, if there are points of

uncertainty, highlight them)

By the end of week 1, we will hope to have contacted radiologists to ask them for help with the evaluating the project, and have prepared most of the datasets. We should be in the middle of writing a GAN model.
By the end of week 2, we will hope to have a working GAN model.
By the end of week 3, we will hope to have applied the GAN model to our entire dataset, and tweaked it for the data. We shall hopefully have images to send to the radiologists
By the end of week 4, we hope to have finished our group poster.

Agreed frequency and mode of communication:
(communicate as much as you can, we also recommend regular check-up meetings in addition to ad hoc communication)

We aim to meet at least two times per week: one on campus during the I-Explore sessions, and another mid-week meeting on campus or through Microsoft Teams. We will actively seek help from Ms. Muller. If anything goes wrong, we will increase the frequency of communication to three or more times a week.

Project repository:
(URL to project's GitHub repository)
https://github.com/arbirdi/IRC_Generative_Medical_imaging

References/websites/AOB:
(sources, websites or anything else you would like to mention)

[1] Singh H. *Everything you Need to Know About Hardware Requirements for Machine Learning*. Available from: https://www.einfochips.com/blog/everything-you-need-to-know-about-hardware-requirements-for-machine-learning/ [Accessed 15th February 2023]

[2] Kim, S., Kim, B. & Park, H. (2021) Synthesis of brain tumor multicontrast MR images for improved data augmentation. *Medical Physics (Lancaster).* 48 (5), 2185–2198. 10.1002/mp.14701.

[3] Yi, X., Walia, E. & Babyn, P. (2019) Generative adversarial network in medical imaging: A review. *Medical Image Analysis.* 58 101552. 10.1016/j.media.2019.101552.

[4] Nasr Esfahani, S. & Latifi, S. (2019) Image Generation with Gans-based Techniques: A Survey. *International Journal of Computer Science and Information Technology.* 11 (5), 33–50. 10.5121/ijcsit.2019.11503.

[5] Simard, P. Y., Steinkraus, D. & Platt, J. C. (2003) Best practices for convolutional neural networks applied to visual document analysis. *ICDAR.* , IEEE. Pp.958–963https://ieeexplore.ieee.org/document/1227801.